

Omar Sherif
Abdallah
Data Analyst



About Me

I've always believed that every dataset has a hidden story waiting to be told. While most people see a wall of confusing numbers, I see a map that points toward better business decisions.

As a Data Analyst, I specialize in translating complex technical findings into clear, actionable insights that even non-technical stakeholders can use. Whether it's cleaning messy data or building interactive dashboards, my goal is always the same: to turn 'what happened' into 'what we should do next'



Education

Academic Background

Bachelor of Business Information SystemsBadr University in Cairo (BUC) | 2022 – June 2026 (Expected)

Courses

"My journey into data analysis is defined by a commitment to continuous learning and practical application. On December 21, 2024, I earned the **Google Data Analytics Professional Certificate** via Coursera, where I mastered the full data lifecycle—from rigorous data cleaning to high-impact visualization using **SQL**, **R**, and **Tableau**. Building on this foundation, I am currently a **Data Analyst** Intern at the **Digital Egypt Pioneers Initiative (DEPI)** in Cairo (Nov 2025 – Present). In this intensive environment, I have significantly expanded my technical proficiency, mastering **Microsoft Excel** for complex data manipulation and **Power BI** for creating interactive, business-ready dashboards. I collaborate on cross-functional teams to solve real-world problems using **SQL** and **Python**, ensuring my technical insights are paired with strong communication and emotional intelligence. My goal is always to turn 'what happened' into 'what we should do next' by delivering data stories that resonate with any audience.



My Services

Messy data is an expensive blind spot. I specialize in turning operational chaos into a clear competitive advantage by identifying hidden revenue leaks and building real-time systems that eliminate guesswork. My Unique Selling Point is my ability to translate complex data into a direct roadmap for growth. I don't just provide information; I provide the clarity you need to make high-stakes decisions that scale your business.

Project: Sales Performance & Profitability Optimization

The Problem: Sales reached \$59.64M, but profit margins were stalled at a low 10.8% due to hidden operational costs.

The Solution: Built a strategic dashboard to track the gap between unit costs and pricing while analyzing regional rep performance.

The Discovery: Found that the business relies 100% on expensive new customer acquisition while ignoring repeat loyalty.

The Recommendation: Proposed a 5% pricing adjustment in Electronics and a re-allocation of resources to underperforming regions.

The Bottom Line: Provided the roadmap to shift from "volume-based" sales to "profit-first" growth.

Sales Performance Dashboard

Payment_Method

Bank Transfer

Cash

Credit Card

59.64M

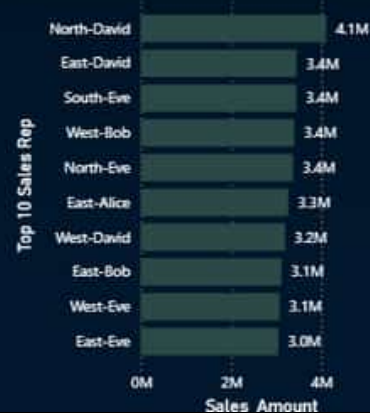
Sum of Sales_Amount

6.47M

Sum of total_profit

25K

Sum of Quantity_Sold



Average of Unit_Cost and Average of Unit_Price by Product_Category



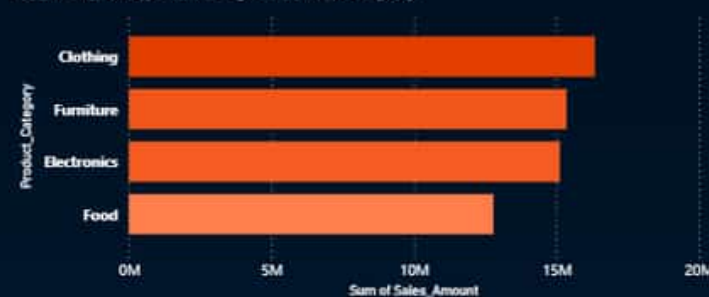
Sum of Sales_Amount by Region



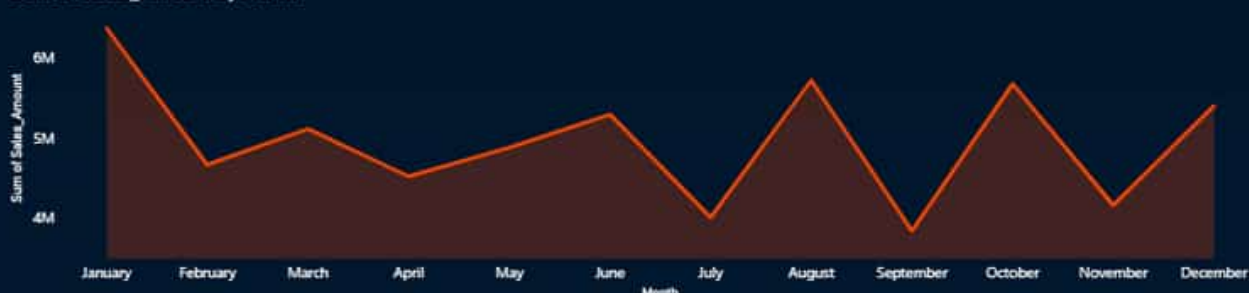
Sum of Sales_Amount by Customer_Type



Sum of Sales_Amount by Product_Category



Sum of Sales_Amount by Month



Key Findings

Sales Distribution: North is the Best Region, but all four regions are very close in total sales, indicating balanced geographical coverage.

Profitability: Profit Margin is Low. Total profit (6.47M) is only about 10.8% of Total Sales (59.64M), suggesting high costs or heavy discounting.

Customer Behavior: New Customers generate significantly more revenue than Returning Customers. **New Product Performance: Furniture and Clothing** are the top-selling product categories by revenue, while Food is the lowest.

Root Causes

Low Profit Margin (10.8%): The **Unit Cost** is very high relative to the **Unit Price** across most categories, or there is significant discounting.

"New" Customers Dominate: The business relies heavily on **acquisition** and is failing to drive repeat purchases, or the loyalty/retention programs are ineffective.

East Region Underperforming: The East region may have a disproportionately high number of low-performing Sales Reps or a low number of total reps.

Recommendations

Low Profit Margin: Review pricing strategy and current discount structure. Target a 5% increase in **Unit Price** or a 5% reduction in **Unit Cost** for the Electronics category where margins look thinnest. **Poor Customer Retention:** Implement a customer loyalty program or targeted re-engagement campaign for 'Returning' customers to boost their average order value and frequency.

Underperforming East Region: Re-allocate sales resources or implement targeted training and coaching for the East region's Sales Reps to bring their average performance in line with the West/South regions.

Project: LA Crime Patterns & Public Safety Intelligence

Using Python (Pandas, Matplotlib, and Seaborn) in a Jupyter environment, I engineered a data pipeline to clean and categorize four years of crime data.

Data Cleaning: Handled missing values across 27 columns and standardized temporal data to extract months, years, and hours.

Trend Analysis: Built time-series models to track how crime volume shifted from 2020 through 2023.

Demographic Profiling: Visualized victim age distributions and gender demographics to identify the most "at-risk" groups.

The Insights (The Strategic Value)

I translated the code into actionable public safety findings:

Peak Risk Hours: Identified a significant spike in criminal activity during the mid-day period (12:00 PM) and late evening, suggesting a need for increased patrol visibility during these windows.

The "Top 3" Threats: Proved that Vehicle Theft is the leading crime, followed by Simple Assault and Identity Theft, allowing for targeted community awareness campaigns.

Age Vulnerability: Discovered that individuals in the 20–40 age bracket are disproportionately targeted, providing a specific demographic for safety education.

Anaconda Toolbox

v4.26.1

Update Available

Update then restart your JupyterLab.

Update Now

Projects

Manage Projects

Code Snippets

Manage Code Snippets

Environments

Create a New Environment

AI Assistant

Chat-based Python Help

jupyter

LA Crime Project Last Checkpoint: 7 months ago

File Edit View Run Kernel Settings Help

Python 3 (pykernel)

Trusted

Load Data

[1]:
import pandas as pd
import matplotlib.pyplot as plt
crime = pd.read_csv('Crime_Data_from_2020_to_Present.csv')

Understand and Clean Data

[2]:
crime.head()

	division_number	date_reported	date_occurred	area	area_name	reporting_district	part	crime_code	crime_description	modus_operandi	
0	10304468	2020-01-08	2020-01-08 22:30:00	3	Southwest	377	2	624	BATTERY - SIMPLE ASSAULT	0444 0913	...
1	190101066	2020-01-02	2020-01-01 03:30:00	1	Central	163	2	624	BATTERY - SIMPLE ASSAULT	0416 1822 1414	...
2	200110444	2020-04-14	2020-02-13 12:00:00	1	Central	155	2	845	SEX OFFENDER REGISTRANT OUT OF COMPLIANCE		...
3	191501505	2020-01-01	2020-01-01 17:30:00	15	Hollywood	1543	2	745	VANDALISM - MISDEAMEANOR (E399 OR UNDER)	0329 1402	...
4	191921269	2020-01-01	2020-01-01 04:15:00	19	Mission	1998	2	740	VANDALISM - FELONY (E400 & OVER, ALL CHURCH VAL...	0329	...

5 rows x 27 columns

[3]:
crime.columns

[3]: Index(['division number', 'date reported', 'date occurred', 'area',

Anaconda Assistant

AI-powered coding, insights and debugging in your notebooks.

To enable the following extensions, create an account or sign in.

Anaconda Assistant 4.26.1

Create Account

Already have an account? [Sign In](#)

For more information, read our [Anaconda Assistant documentation](#).

Anaconda Toolbox

v4.26.1

Update Available

Update then restart your JupyterLab.

Update Now

Projects

Manage Projects

Code Snippets

Manage Code Snippets

Environments

Create a New Environment

AI Assistant

Chat-based Python Help

jupyter

LA Crime Project Last Checkpoint: 7 months ago

File Edit View Run Kernel Settings Help

Python 3 (pykernel)

Trusted

[3]:
crime.columns

[3]: Index(['division number', 'date reported', 'date occurred', 'area', 'area_name', 'reporting_district', 'part', 'crime_code', 'crime_description', 'modus_operandi', 'victim_age', 'victim_sex', 'victim_descent', 'premise_code', 'premise_description', 'weapon_code', 'weapon_description', 'status', 'status_description', 'crime_code_1', 'crime_code_2', 'crime_code_3', 'crime_code_4', 'location', 'cross_street', 'latitude', 'longitude'], dtype='object')

[4]:
crime.shape

[4]: (852950, 27)

[5]:
crime.dtypes

[5]:

division_number	int64
date_reported	object
date_occurred	object
area	int64
area_name	object
reporting_district	int64
part	int64
crime_code	int64
crime_description	object
modus_operandi	object
victim_age	int64
victim_sex	object
victim_descent	object
premise_code	float64
premise_description	object
weapon_code	float64
weapon_description	object
status	object
status_description	object
crime_code_1	float64
crime_code_2	float64
crime_code_3	float64
crime_code_4	float64
location	object

Anaconda Assistant

AI-powered coding, insights and debugging in your notebooks.

To enable the following extensions, create an account or sign in.

Anaconda Assistant 4.26.1

Create Account

Already have an account? [Sign In](#)

For more information, read our [Anaconda Assistant documentation](#).

Anaconda Toolbox

Update Available

Update then restart your JupyterLab.

Update Now

Projects

Manage Projects

Code Snippets

Manage Code Snippets

Environments

Create a New Environment

AI Assistant

Chat-based Python Help

Jupyter

LA Crime Project

Last Checkpoint: 7 months ago

File

Edit

View

Run

Kernel

Settings

Help

JupyterLab

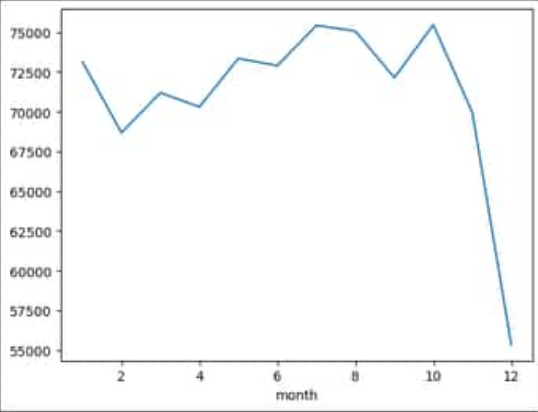
Python 3 (pykernel)

Anaconda Toolbox

Feature Understanding

[18]: df['month'].value_counts().sort_index().plot()

[18]: <Axes: xlabel='month'>



[19]: df['year'].value_counts().sort_index().plot(kind='bar')

[19]: <Axes: xlabel='year'>

Anaconda Assistant

AI-powered coding, insights and debugging in your notebooks.

To enable the following extensions, create an account or sign in.

Anaconda Assistant

4.26.1

Create Account

Already have an account? [Sign In](#)

For more information, read our [Anaconda Assistant documentation](#).

Anaconda Toolbox

Update Available

Update then restart your JupyterLab.

Update Now

Projects

Manage Projects

Code Snippets

Manage Code Snippets

Environments

Create a New Environment

AI Assistant

Chat-based Python Help

Jupyter

LA Crime Project

Last Checkpoint: 7 months ago

File

Edit

View

Run

Kernel

Settings

Help

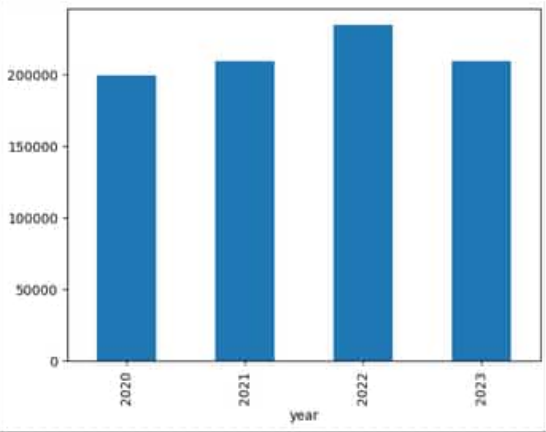
JupyterLab

Python 3 (pykernel)

Anaconda Toolbox

[19]: df['year'].value_counts().sort_index().plot(kind='bar')

[19]: <Axes: xlabel='year'>



[20]: df = df[~((df['month'] == 12) & (df['year'] == 2023))]

[21]: df.sort_values(by = 'date_occurred')

[21]:

	division_number	date_reported	date_occurred	area	area_name	reporting_district	part	crime_code	crime_description	victim_age	...
168264	201810982	2020-05-11	2020-01-01 00:01:00	18	Southeast	1822	2	810	SEXUALUNLAWFULINCMUTUAL CONSENT, PENETRATION W...	17	-
196441	220706832	2022-03-16	2020-01-01 00:01:00	7	Wiltshire	782	2	668	EMBEZZLEMENT, GRAND THEFT (\$950.01 & OVER)	0	-

Anaconda Assistant

AI-powered coding, insights and debugging in your notebooks.

To enable the following extensions, create an account or sign in.

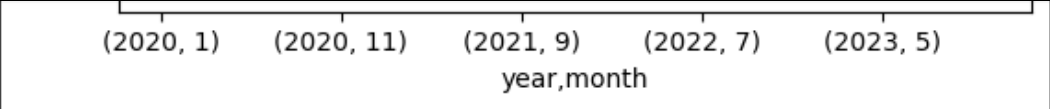
Anaconda Assistant

4.26.1

Create Account

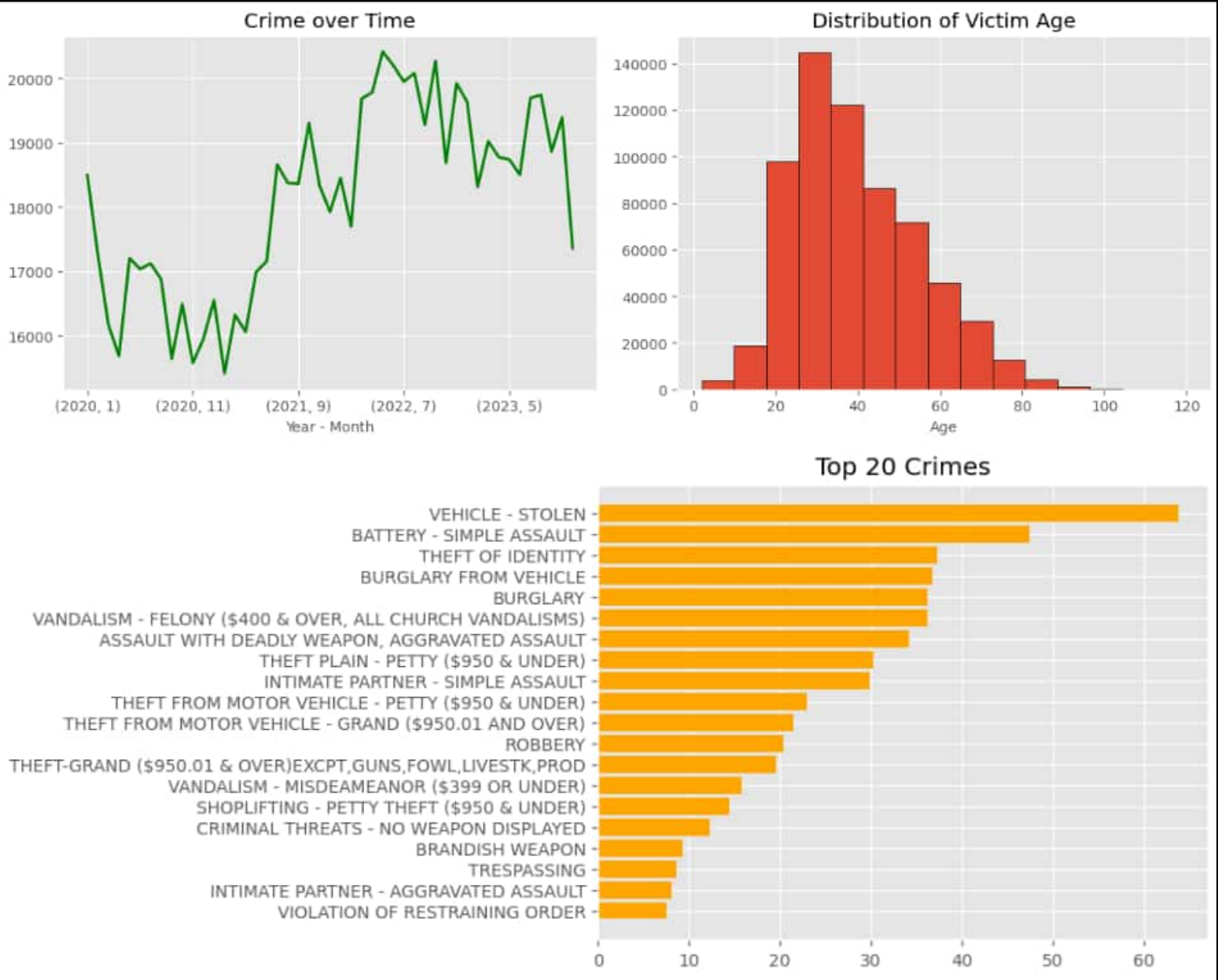
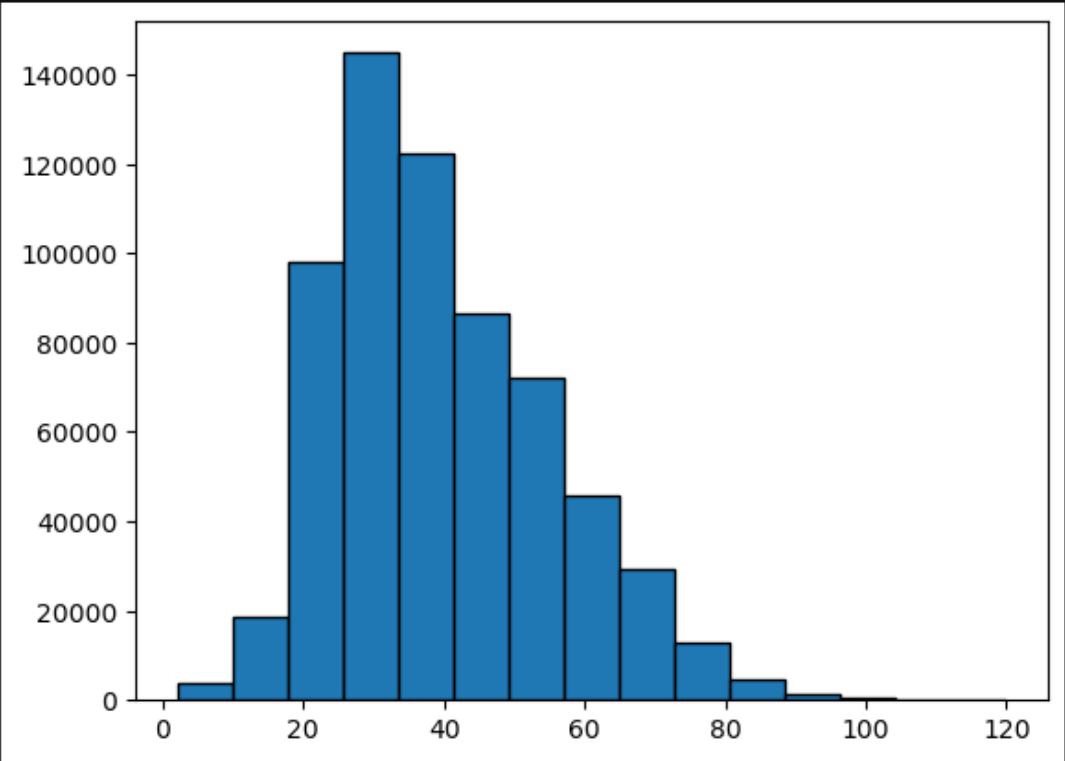
Already have an account? [Sign In](#)

For more information, read our [Anaconda Assistant documentation](#).



```
[23]: plt.hist(df_age['victim_age'], edgecolor = 'black', bins = 15)
```

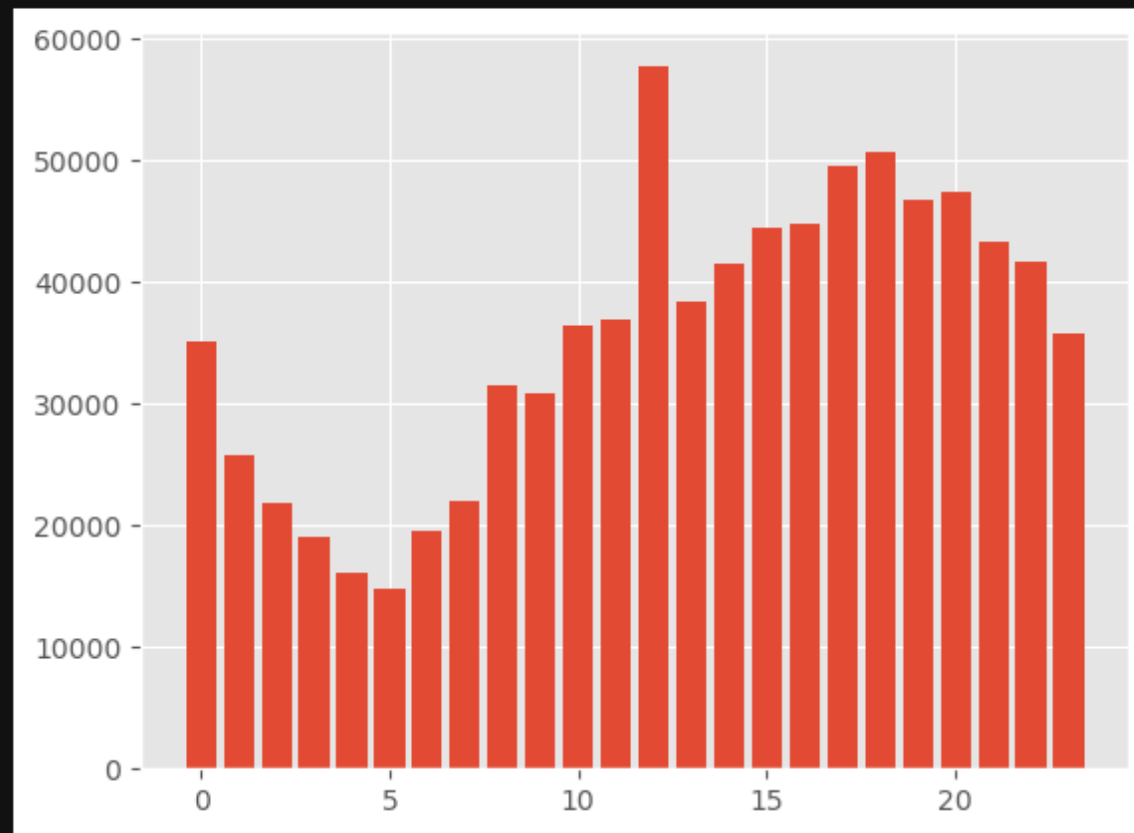
```
[23]: (array([3.96800e+03, 1.86630e+04, 9.81120e+04, 1.44858e+05, 1.22469e+05,
        8.64850e+04, 7.20760e+04, 4.59690e+04, 2.94150e+04, 1.28000e+04,
        4.50300e+03, 1.27400e+03, 4.41000e+02, 0.00000e+00, 1.00000e+00]),
      array([ 2.          ,  9.86666667, 17.73333333, 25.6          ,
        33.46666667, 41.33333333, 49.2          , 57.06666667,
        64.93333333, 72.8          , 80.66666667, 88.53333333,
        96.4          , 104.26666667, 112.13333333, 120.          ]),
      <BarContainer object of 15 artists>)
```



```
[40]: hourly_crime = df['date_occurred'].dt.hour.value_counts().sort_index().reset_index()
```

```
[41]: plt.bar(hourly_crime['date_occurred'], hourly_crime['count'])
```

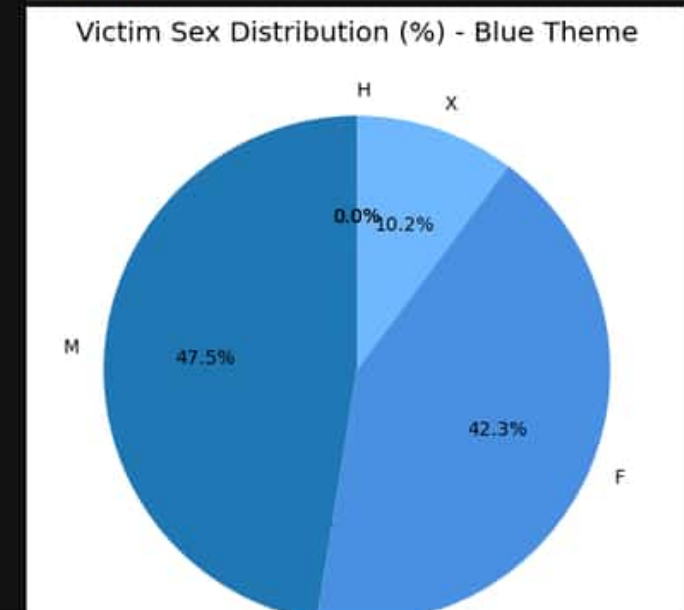
```
[41]: <BarContainer object of 24 artists>
```



```
[47]: # Calculate percentages
gender_pct = df['victim_sex'].value_counts(normalize=True)

# Blue color palette
colors = ['#1f77b4', '#4a9be2', '#6fb8ff', '#9ecbff', '#cfe4ff']

plt.figure(figsize=(6,6))
plt.pie(
    gender_pct,
    labels=gender_pct.index,
    colors=colors[:len(gender_pct)], # Match number of categories
    autopct='%1.1f%%',
    startangle=90
)
plt.title('Victim Sex Distribution (%) - Blue Theme')
plt.show()
```



Project: HR Workforce & Compensation Intelligence

The Problem: The organization lacked a centralized view of its 1,000 employees, making it difficult to track compensation fairness, department-level spending, or hiring trends across its five regional branches.

The Solution: I developed a multi-platform analysis using Excel for complex data structuring and a custom Interactive Dashboard for executive visibility.

Data Architecture: Organized 20 years of hiring data to calculate an average of 20.0 years of experience per employee.

Compensation Modeling: Built pivot tables to analyze a total payroll of \$9.3M, breaking down average salaries (\$9,326) by department, gender, and pay type.

The Insights:

Operational Transparency: Identified that Compliance and Maintenance are the highest-cost departments, while Logistics remains the leanest.

Equity Audit: Visualized the salary gap between Managers and Supervisors to ensure the compensation structure remains competitive and fair.

Branch Efficiency: Discovered that the Sharqia branch holds the highest employee count, providing a roadmap for future resource re-allocation.

Employees Analysis Dashboard

TOTAL Employees

Average Salaries

Average Years Of Experience

1000

\$9,326.24

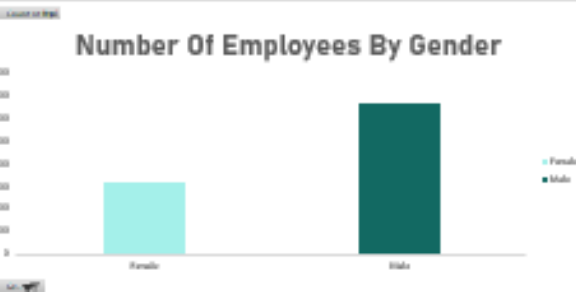
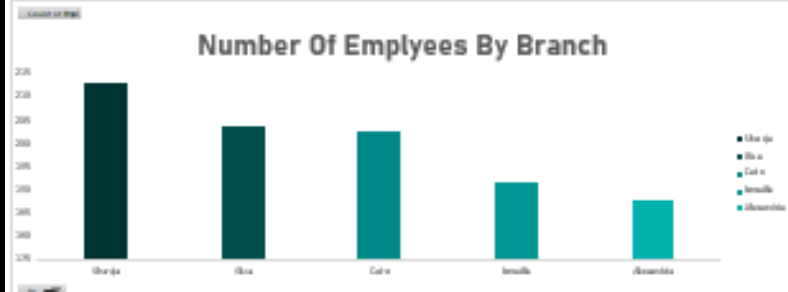
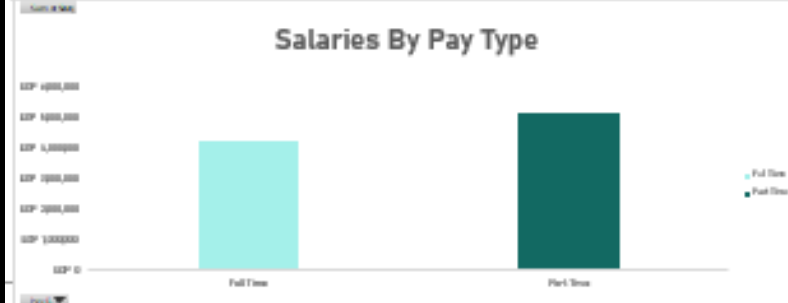
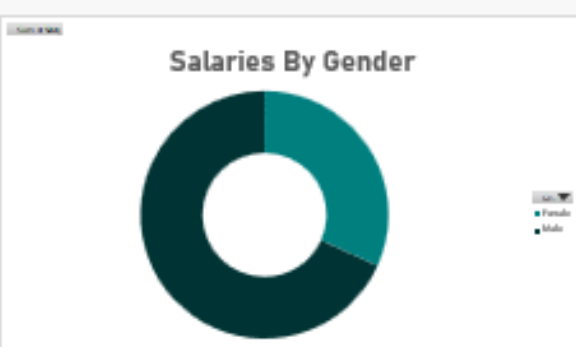
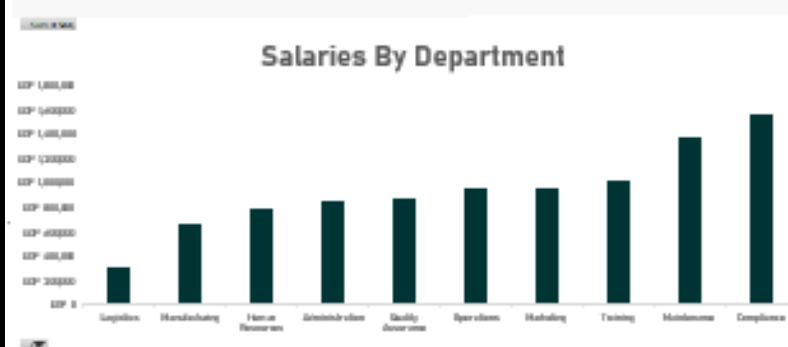
20.0

PayType
Full Time
Part Time

Branch
Alexandria
Cairo
Giza
Ismailia
Sharqia

Gender
Female
Male

Position
Employee
Manager
Supervisor



AutoSave Off

hr analytics cls (version 2).xlsx

Search

FileHomeInsertPage LayoutFormulasDataReviewViewAutomateHelp

Clipboard

Font

Alignment

Number

Styles

Cells

Editing

Add-ins

M24

Departments	Sum of Salary	Gender	Count of EmplID	Gender	Sum of Salary
Logistics	EGP 310,846	Female	326	Female	EGP 2,961,855
Manufacturing	EGP 654,860	Male	674	Male	EGP 6,364,383
Human Resources	EGP 797,470	Grand Tot	1000	Grand Total	EGP 9,326,238
Administration	EGP 848,561				
Quality Assurance	EGP 861,686				
Operations	EGP 947,425				
Marketing	EGP 953,198				
Training	EGP 1,009,969				
Maintenance	EGP 1,371,364				
Compliance	EGP 1,570,859				
Grand Total	EGP 9,326,238				
Hire Date In Years	Count of EmplID	Branch	Sum of Salary	Branch	Count of EmplID
1995	44	Alexandria	EGP 1,878,121	Alexandria	188
1996	38	Cairo	EGP 1,941,358	Cairo	203
1997	52	Giza	EGP 1,964,781	Giza	204
1998	52	Ismailia	EGP 1,744,676	Ismailia	192
1999	48	Sharqia	EGP 1,797,302	Sharqia	213
2000	31	Grand Total	EGP 9,326,238	Grand Total	1000
2001	50				
2002	45				
2003	58				
2004	52				
2005	44				
2006	50				
2007	47				
2008	56				
2009	41				
2010	59				
2011	57				
2012	45				
2013	41				
2014	45				
2015	45				
Grand Total	1000				
Position	Sum of Salary	Position	Sum of Salary		
Manager	EGP 3,954,662	Manager	EGP 3,954,662		
Employee	EGP 3,359,655	Employee	EGP 3,359,655		
Supervisor	EGP 2,011,921	Supervisor	EGP 2,011,921		
Grand Total	EGP 9,326,238	Grand Total	EGP 9,326,238		

Bank Misr Performance Analysis

Segment Concentration: Identified that VIP clients hold 70.25% of total balances, outperforming the Retail segment by 670% in value despite a smaller headcount.

Portfolio Health: Verified a robust operational status with 65.17% (391 loans) active in the current portfolio.

Market Shifts: Analyzed a strategic pivot where Car Loans grew by 8.94%, successfully offsetting a 72.22% contraction in Mortgages amid a general 20% decline in loan volume since 2018.

Customer Demographics: Determined that Retirees represent the largest customer block (134 clients), followed by high-value professional groups (Engineers and Doctors).

Operational Trends: Observed a 4.38% increase in 2024 transaction volume, driven primarily by Deposits, signaling high customer trust and liquidity.

Geographic Footprint: Confirmed centralist expansion patterns, with 25% of the total branch network concentrated in Cairo.



ACCOUNTS

BRANCHES

CUSTOMERS

LOANS

Transactions

CLEAR FILTERS

Accounts Analysis

Branch_Name

All



1.243K
Total Accounts

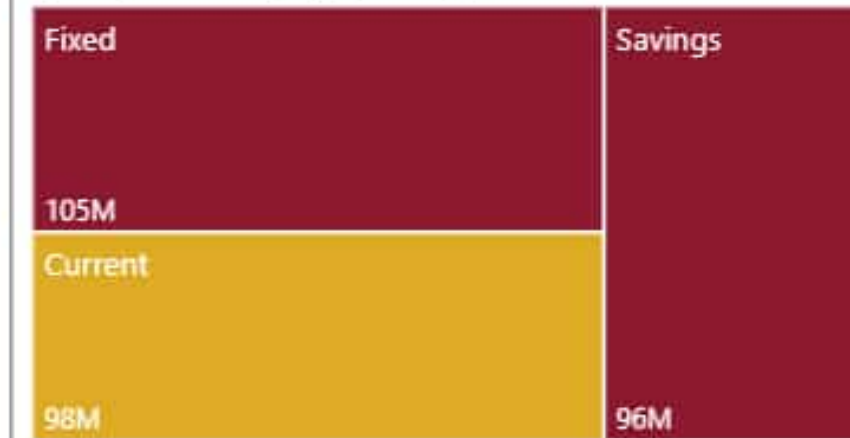


299M
Total Balance

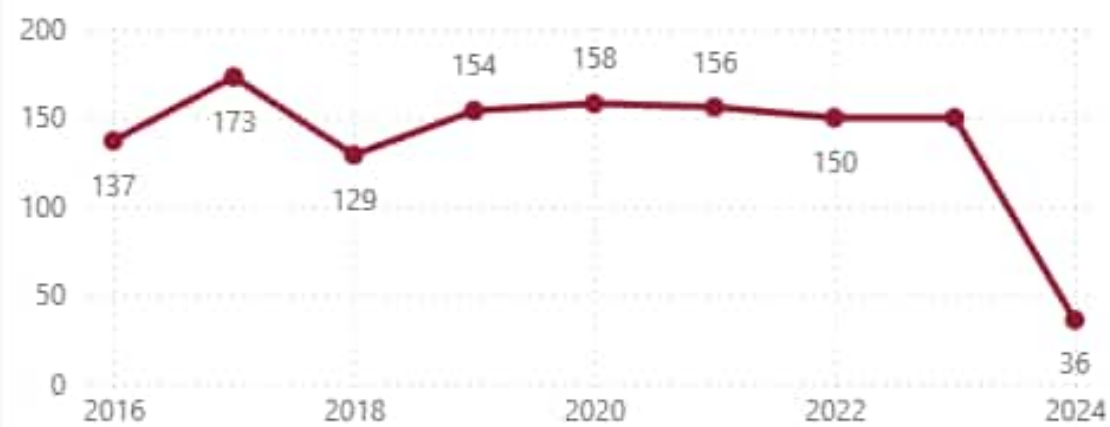
Account Portfolio Breakdown by Type and Segment

Segment	Current	Fixed	Savings	Total
Mass	6.89	7.15	7.03	7.02
Retail	6.96	6.98	6.94	6.96
VIP	6.67	6.93	6.68	6.77
Total	6.91	7.02	6.94	6.96

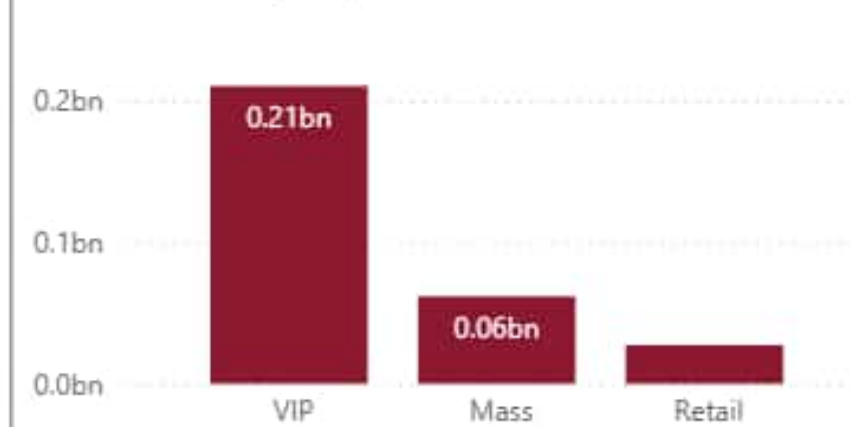
Total Balance by Account Type



Total Accounts by Year



Total Balace by Segment



Project: BMW Strategic Performance Analysis (2015–2024)

- **Financial Growth:** Analyzed \$48bn in revenue and 633K units sold, maintaining a steady 12.1% annual growth.
- **The Electric Shift:** Tracked BMW's pivot to EVs, which now lead the portfolio with 34.7% of total revenue (\$17bn).
- **Top Performers:** Identified China as the top market (\$6.1bn) and the X3 as the highest-grossing model.
- **Marketing Efficiency:** Achieved a high Return on Ad Spend (172x) while keeping the marketing cost per car low at \$441.
- **Strategic Allocation:** Linked the \$16M spend on the X3 to its massive \$3.16bn revenue outcome.



OVERVIEW

Product & Engine
Analysis

Marketing &
Market Strategy

CLEAR FILTER

OVERVIEW

Drill Down Period

All

\$48bn

Total Revenue

633K

Total Units Sold

10

Total Models

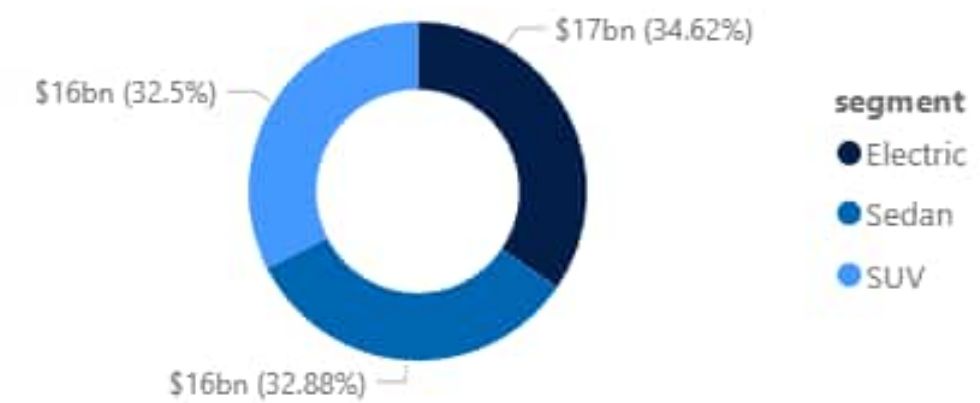
12.12%

Unit Growth (YOY)

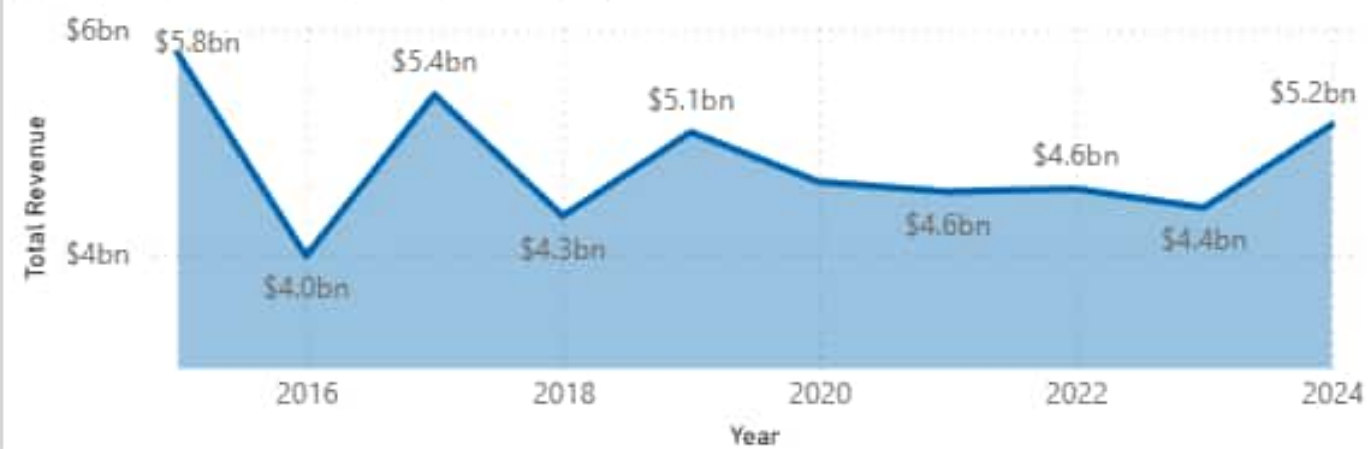
Global Revenue Distribution



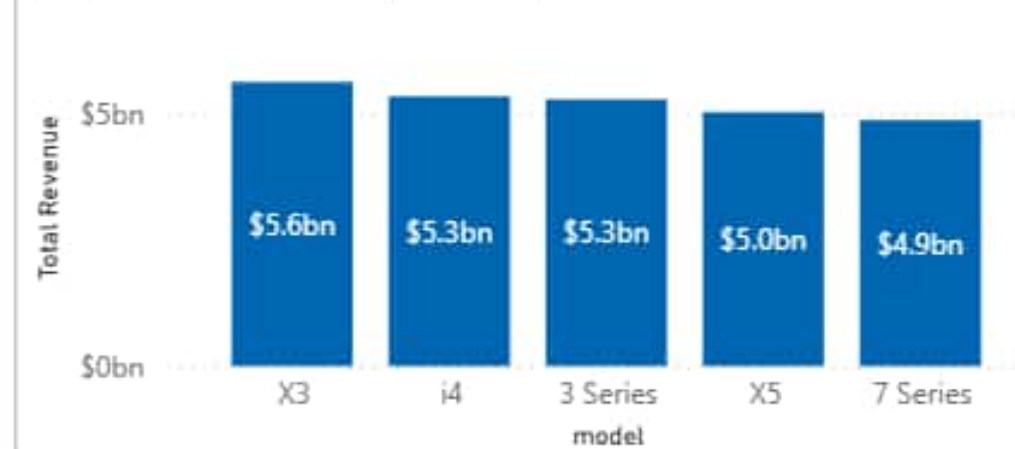
Revenue Share by Vehicle Category



Annual Revenue Performance Trend



Top 5 Revenue-Generating Models



Project: Video Games Historical Sales Analysis

Data Preparation: Processed a dataset of 64K records, filtering it down to 17.5K active entries by removing zero-sales "noise" for higher accuracy.

Data Cleaning: Handled missing values using Median Imputation for Critic Scores (by Genre) and Release Years (by Platform) to ensure data consistency.

Global Impact: Analyzed 6.6B total copies sold across 39 platforms, with 2008 marking the all-time sales peak (538.8M units).

Market Leaders: Identified PS2 as the top-selling platform (1.03B units) and Sports as the most popular genre (1.19B units).

Regional Dominance: Found that North America (NA) controls over half the global market share at 51.12%.

Top Performers: Highlighted GTA V as the best-selling title (64M copies) and Activision as the leading publisher (772.6M units).

Video Games Analysis

6.60bn

Sum of Total Units...

12K

Total Games

39

Total Consoles

Release Date

1/1/1977

12/31/2020

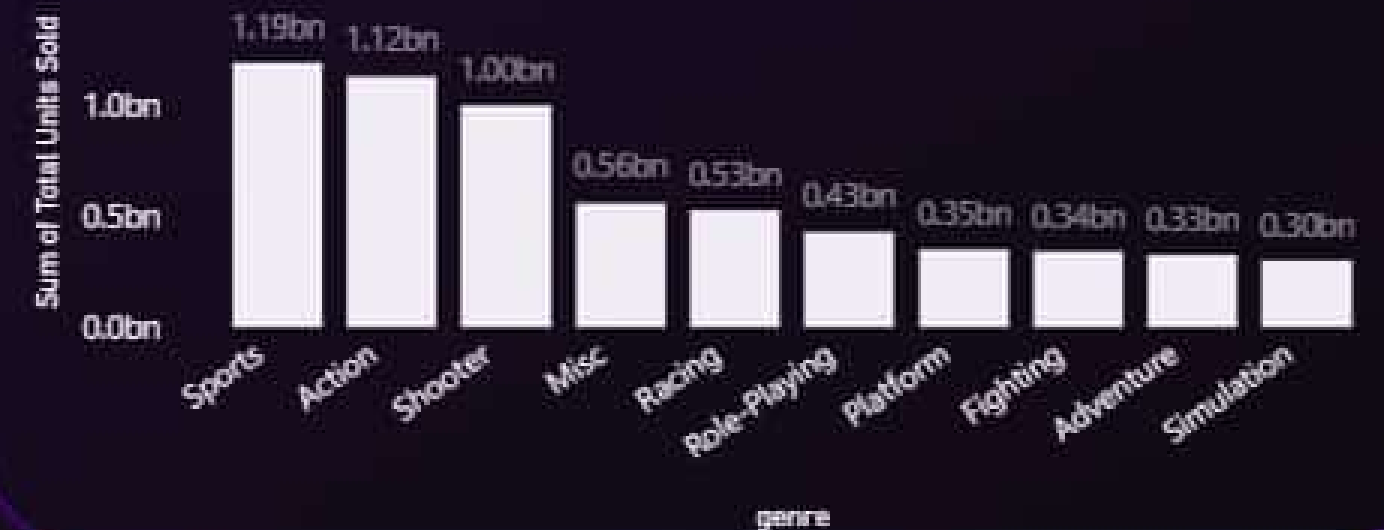
Games

All

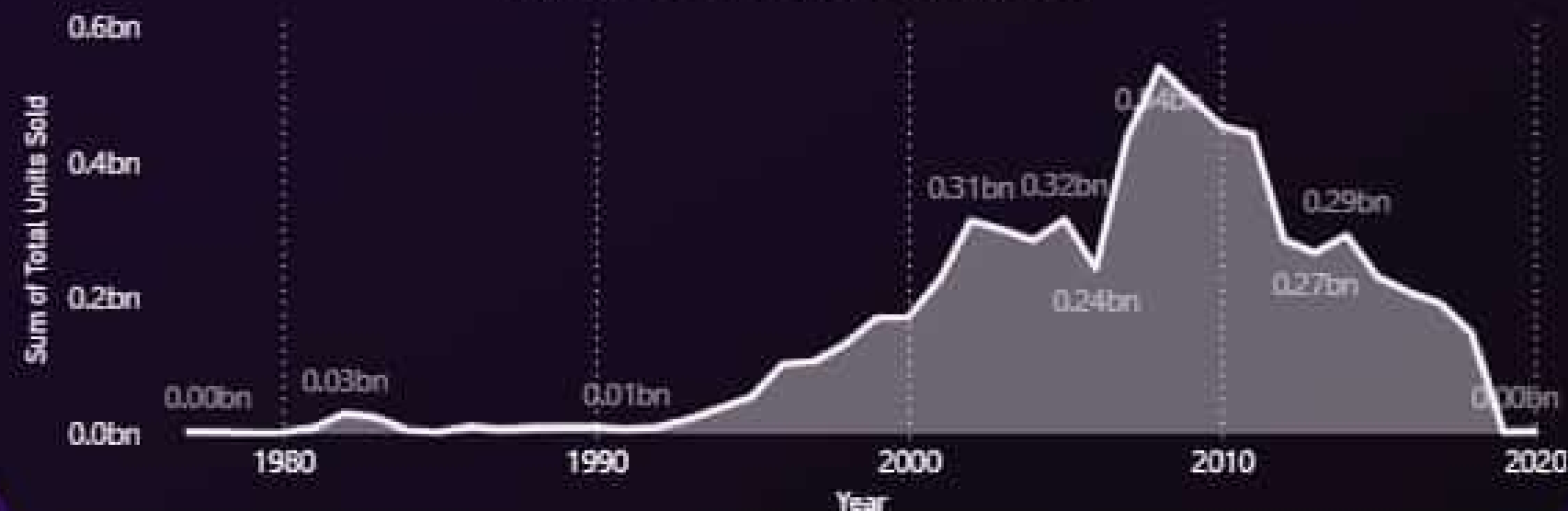
Top 10 Consoles By Units Sold



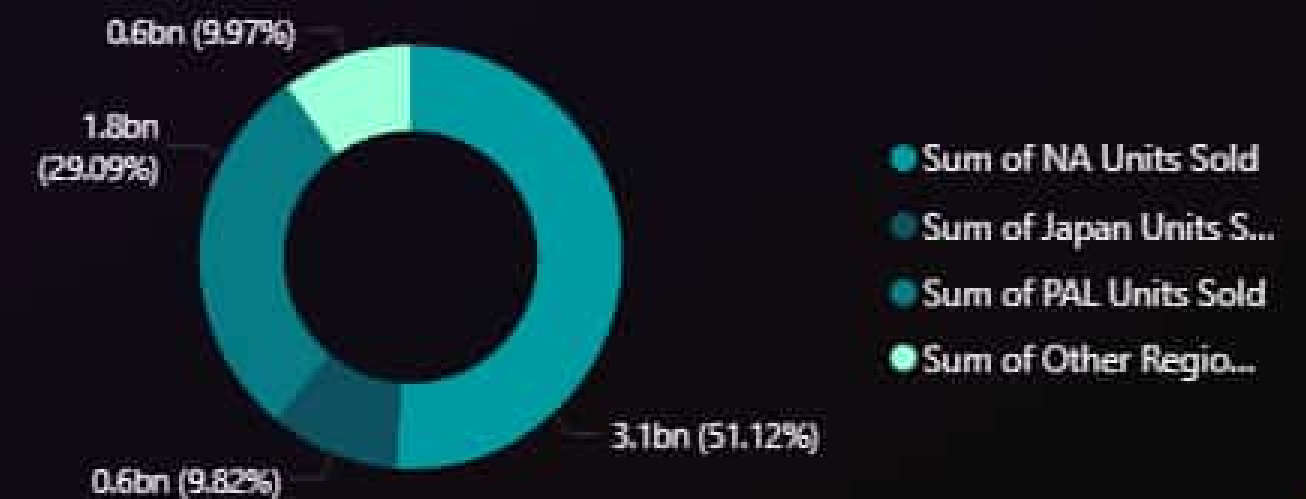
Top 10 Genre By Units Sold



Sum of Total Units Sold by Year



Units Sold By Region



Get in Touch

Phone Number

01002868726

Email

Omar.s.sheshtawy@gmail.com

linkedin:

<http://www.linkedin.com/in/omar-sherif-540368255>

